

PROJETO FINAL - CHECKPOINT 2: ESCOLHA DO MODELO

ADALBERTO FERREIRA BARBOSA JUNIOR
JULIANA RESPLANDE SANT'ANNA GOMES
LARISSA SILVA XAVIER ROSA
RAFAEL NOGUEIRA CUNHA

UFG – Universidade Federal de Goiás
Goiânia - GO

Resumo: Este artigo propõe a identificação e classificação de doenças de pele pelo modelo Vision Transformers, utilizando dados de imagens provenientes do Dataset “Dermnet Skin Disease Images”, disponível no Kaggle. Os resultados mostram que o modelo atingiu acurácia de 85% nos dados de teste, provando que a arquitetura Transformer originalmente desenvolvida para tarefas de processamento de linguagem natural é eficiente para tarefas de classificação em visão computacional.

Palavras Chaves: Doenças de Pele, Vision Transformers, ViT, Imagens, Classificação.

1. INTRODUÇÃO

No que se trata do âmbito da Saúde, as doenças de Pele geralmente são tratadas com menos preocupação, por terem menor letalidade quando comparadas à outras. No entanto, esse tipo de doenças atinge uma grande porção de indivíduos, o que faz necessária uma maior atenção a elas.

A pele é o maior órgão do corpo e possui diversas funções, como manter os fluidos do organismo, proteger contra bactérias, vírus e agressões externas, regular a temperatura corporal, entre outras.

Nesse sentido, deve-se considerar as possíveis perturbações da pele, como irritações, obstruções ou inflamações como fatores de observação dessas doenças, que causam desde sintomas leves como vermelhidão, inchaço, queimação e coceira, até erupções cutâneas, urticária e outros problemas de pele. Muitas doenças de pele, como a acne, afetam a aparência, além de prejudicar a qualidade de vida do paciente.

A maioria das desordens e doenças da pele são facilmente tratáveis, mas algumas são

potencialmente graves. Existem centenas de doenças de pele, para citar algumas: celulite, estrias, acne, verruga, dermatite, eczema, sarna, urticária entre outras.

2. METODOLOGIA

2.1 Dados

Como estudo de caso foi utilizado o dataset “Dermnet Skin Disease Images”, disponível no Kaggle [1]. O conjunto de dados compreende 23 categorias de doenças de pele, em que as amostras já estão divididas inicialmente entre treino e teste.

2.1.1 Classes de dados utilizadas

O conjunto de dados contém 23 classes com cada classe variando consideravelmente no número de exemplos: a classe majoritária contém cerca de 2,5 mil exemplos, enquanto a classe minoritária possui 330 exemplos. Para obter uma distribuição homogênea dos dados, foram consideradas inicialmente as 4 Classes Majoritárias, isto é, as

que mais possuem números de exemplos. No entanto, trocou-se a 3ª classe majoritária que representa a patologia Molusco pela patologia de acne.

	class	n_samples
Psoriasis pictures Lichen Planus and related d...		2501
Actinic Keratosis Basal Cell Carcinoma and oth...		2139
Seborrheic Keratoses and other Benign Tumors		2010
Acne and Rosacea Photos		1779

Tabela 1

Essa escolha foi feita, pelo fato de Molusco Infeccioso ser uma doença infecciosa atípica relativamente comum somente às crianças, enquanto a acne é mais frequente na população em diversas faixas etárias.

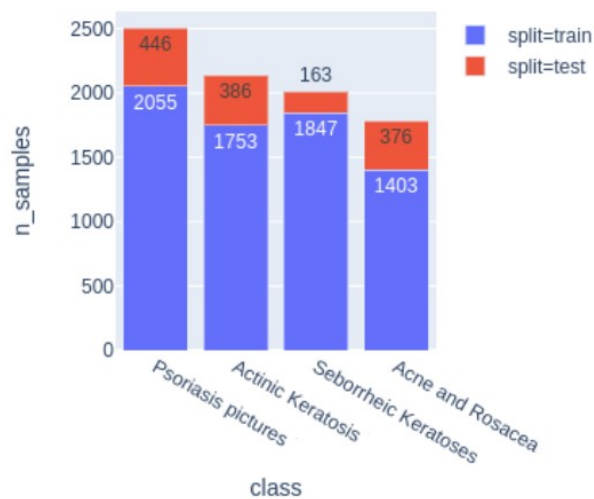


Figura 1

A Tabela 1 indica as classes escolhidas, em que *class* representa o nome da classe e *n_samples* o número de exemplos do conjunto de treino e de teste. Essa análise das classes é minuciada na Figura 1, que também quantifica o número de exemplos entre a divisão de dados de treino e de teste. Em seguida, há uma breve descrição de cada classe.

Psoriasis pictures Lichen Planus and related diseases:



Figura 2

Doença da pele comum, crônica e não contagiosa. É cíclica, ou seja, apresenta sintomas que desaparecem e reaparecem periodicamente, um exemplo da doença pode ser visualizado na Figura 2. Sua causa é desconhecida, mas se sabe que pode estar relacionada ao sistema imunológico, às interações com o meio ambiente e à suscetibilidade genética. Devido ao seu ciclo de ataques e grande produção de células, há a formação de manchas espessas e escamosas na pele. Normalmente, essa cadeia só é quebrada com tratamento.

Actinic Keratosis Basal Cell Carcinoma and other Malignant Lesions:



Figura 3

É uma alteração da camada mais superficial da pele (camada córnea), com hipertrofia e aspecto escamoso ou verrucoso, sendo que o número de lesões pode ser variado. A Figura 3 ilustra um exemplo do conjunto de treino.

Ceratoses actínicas são lesões pré-cancerosas dos queratinócitos e são muito comuns na população idosa. Eles geralmente ocorrem em áreas expostas ao sol, como as áreas da cabeça e pescoço, bem como as extremidades distais. A exposição à luz ultravioleta (UV) é sua causa mais comum, sendo também importante uma predisposição genética. Por ser uma lesão precursora do Carcinoma Espinocelular (o tipo mais comum de câncer de pele), deve ser tratada com terapia adequada.

Acne and Rosacea:



Figura 4

É um processo inflamatório das glândulas sebáceas e dos folículos pilossebáceos, causando o surgimento de cravos e espinhas, comum tanto na adolescência quanto na fase adulta. A Figura 4 ilustra um exemplo do conjunto de treino. As lesões aparecem com mais frequência na face, mas também podem ocorrer nas costas, ombros e peito.

Além do incômodo das lesões, como na adolescência a aparência é um fator importante, o comprometimento estético determinado por alterações da pele pode atingir o lado psicológico e tornar o adolescente inseguro, com consequências psicológicas que podem persistir pelo resto da vida. Somado a isso, tem-se ainda a aparição de cicatrizes e manchas na pele. No entanto, a acne não é contagiosa e não se relaciona à “sujeira” da pele ou do sangue.

Seborrheic Keratoses and other Benign Tumors:



Figura 5

É uma lesão benigna da pele, geralmente arredondada ou irregular, de coloração acastanhada, amarronzada ou negra, e de aspecto verrucoso. Aparece principalmente na face e tronco e pode crescer se tornando volumosa, geralmente é de origem genética. A Figura 5 ilustra uma amostra do conjunto de dados.

2.1.2 Divisão dos dados

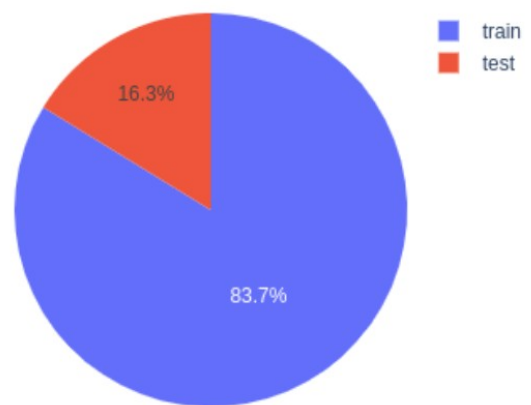


Figura 5: Divisão dos dados

Removendo as classes minoritárias, os dados foram divididos entre 83,7% para treino e 16,3% para teste, resultado obtido da divisão já fornecida no conjunto de dados, a Figura 5 ilustra essa divisão com um gráfico de setores.

2.2. Vision Transformers

Vision Transformers (ViT) é um modelo de aprendizado supervisionado para processamento de visão, originalmente inspirado em modelos baseados em texto (Transformers). Diferente de modelos tradicionais para classificação de imagens, os ViT não possuem redes neurais convolucionais (CNN) na sua arquitetura.

O mecanismo principal por trás da arquitetura do Transformer é a Autoatenção [8]. Esse mecanismo é responsável por capacitar o modelo a aprender e identificar conexões entre suas entradas. Quando os Transformers são aplicados para Processamento de Linguagem Natural (NLP), é calculada a relação entre as palavras de maneira bidirecional, o que significa que a ordem de entrada não importa, diferente dos modelos de redes neurais recorrentes (RNN). Um modelo com arquitetura Transformer lida com entradas de tamanho variável usando pilhas de camadas de autoatenção em vez de CNNs e RNNs.

Um grande desafio de aplicar Transformers a imagens sem o uso de CNNs é aplicar Autoatenção entre pixels. Por exemplo, se o tamanho da imagem de entrada é 640 x 640, o modelo precisa calcular a autoatenção para combinações de 409K. Além disso, é pouco provável que um pixel no canto de uma imagem tenha uma relação significativa com outro pixel na extremidade oposta. O ViT superou esse problema segmentando as imagens em pequenos patches (como 16x16).

2.2.1 Ferramentas utilizadas

O código será executado na ferramenta Google Colaboratory, baseada em Jupyter Notebook, o qual facilita as execuções de código em Python. O framework na utilização do modelo é o HuggingFace Transformers, em conjunto com a biblioteca de otimização Flax da Google.

3. RESULTADOS

3.1 Avaliação do Modelo

Durante o treinamento, foi utilizado o conjunto de dados de teste para a validação do modelo, a Figura 6 ilustra o gráfico da precisão do modelo gerado pelo TensorBoard, framework do TensorFlow que fornece visualização de métricas. No gráfico é possível verificar o comportamento da precisão ao longo das iterações do treinamento, a acurácia atingida na validação foi de 0,8467.

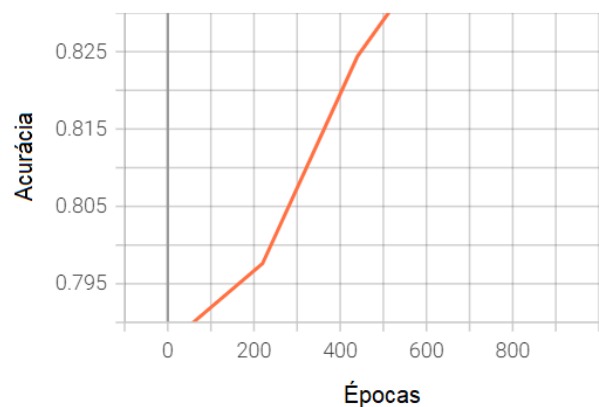


Figura 6: Gráfico da precisão do modelo.

Também foi gerado pelo TensorBoard o gráfico de comportamento da função de custo do modelo durante o treinamento, a qual informa quão bem o modelo performa nos dados de treino e pode ser visualizado na Figura 7. No decorrer das iterações, o modelo atingiu perda igual à 0,4518.

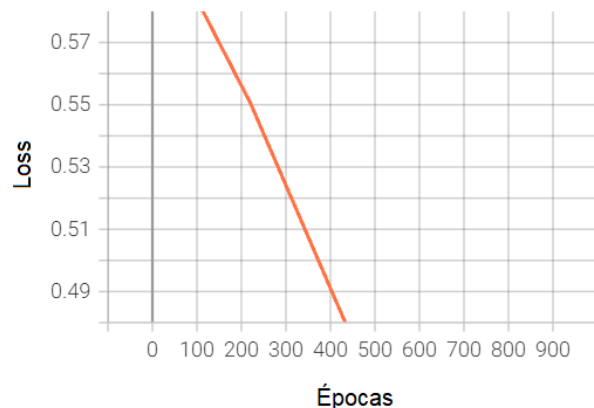


Figura 7: Função de custo.

3.2 Métricas

Como métricas de avaliação, foram utilizadas a matriz de confusão do modelo e o relatório de classificação, ambas métricas fornecidas pela biblioteca de machine learning scikit-learn.

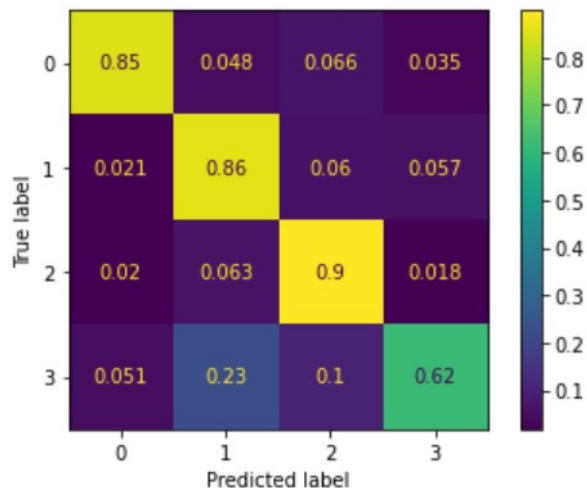


Figura 8: Matriz de confusão normalizada.

A Figura 8 ilustra a matriz de confusão do modelo, já normalizada de acordo com o número de amostras para cada classe. Na diagonal principal, é possível verificar as classes que foram corretamente classificadas de acordo com sua *label*.

Nas três primeiras linhas da matriz, as classes de Acne e Rosácea (0), Ceratose Actínica (1) e Psoríase (2) tiveram, respectivamente, 85%, 86% e 90% das suas amostras classificadas corretamente pelo modelo, enquanto a última classe de Ceratose Seborreica (3) obteve apenas 62% de suas amostras corretamente classificadas. Uma possível causa disso pode ser o baixo número de amostras separadas para teste na classe 4, quando em comparação com as outras classes, como observado no *support* do relatório de classificação (Figura 9).

	precision	recall	f1-score	support
0	0.93	0.85	0.89	376
1	0.81	0.86	0.84	386
2	0.87	0.90	0.88	446
3	0.66	0.62	0.64	136
accuracy			0.85	1344
macro avg	0.82	0.81	0.81	1344
weighted avg	0.85	0.85	0.85	1344

Figura 9: Relatório de Classificação

4. CONCLUSÕES

Segundo os resultados observados, pode-se concluir que o modelo de processamento de linguagem natural Transformers é eficiente para tarefas de visão computacional. No *case* de classificação de doenças de pele, foi atingida acurácia de 85% nos dados de teste para as quatro classes selecionadas.

5. REFERÊNCIAS

- [1]<https://www.kaggle.com/umairshahab/dermnet-skin-disease-images>
- [2]<https://ichi.pro/pt/voce-esta-pronto-para-o-vision-transformer-vit-221968613950777>
- [3]<https://dicassobresaude.com/doencas-de-pele-conheca-as-10-mais-comuns/>
- [4]<https://www.sbd.org.br/dermatologia/pele/doencas-e-problemas/>
- [5]<https://accessmedicine.mhmedical.com/content.aspx?bookId=2184§ionId=165460313>
- [6]<https://accessmedicine.mhmedical.com/content.aspx?sectionid=258956562&bookid=3081&Resultclick=2>
- [7]<https://www.ncbi.nlm.nih.gov/books/NBK11733/>
- [8] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems* (pp. 5998-6008).